


Illumination angle correction during image acquisition in light-sheet fluorescence microscopy using deep learning

CHEN LI,^{1,2} MANI RATNAM RAI,^{1,2} H. TROY GHASHGHAEE,^{2,3} AND ALON GREENBAUM^{1,2,4,*} 

¹Joint Department of Biomedical Engineering, North Carolina State University and University of North Carolina at Chapel Hill, Raleigh, NC 27695, USA

²Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27695, USA

³Department of Molecular Biomedical Sciences, North Carolina State University, Raleigh, NC 27695, USA

⁴Bioinformatics Research Center, North Carolina State University, Raleigh, NC 27695, USA

*greenbaum@ncsu.edu

Abstract: Light-sheet fluorescence microscopy (LSFM) is a high-speed imaging technique that provides optical sectioning with reduced photodamage. LSFM is routinely used in life sciences for live cell imaging and for capturing large volumes of cleared tissues. LSFM has a unique configuration, in which the illumination and detection paths are separated and perpendicular to each other. As such, the image quality, especially at high resolution, largely depends on the degree of overlap between the detection focal plane and the illuminating beam. However, spatial heterogeneity within the sample, curved specimen boundaries, and mismatch of refractive index between tissues and immersion media can refract the well-aligned illumination beam. This refraction can cause extensive blur and non-uniform image quality over the imaged field-of-view. To address these issues, we tested a deep learning-based approach to estimate the angular error of the illumination beam relative to the detection focal plane. The illumination beam was then corrected using a pair of galvo scanners, and the correction significantly improved the image quality across the entire field-of-view. The angular estimation was based on calculating the defocus level on a pixel level within the image using two defocused images. Overall, our study provides a framework that can correct the angle of the light-sheet and improve the overall image quality in high-resolution LSFM 3D image acquisition.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Light-sheet fluorescence microscopy (LSFM) has become an indispensable tool for discovery in the life science community for its high throughput capacity in imaging of live or cleared fixed tissues [1–12]. In LSFM, the sample is illuminated by a thin sheet of light, and fluorophores inside the confined excitation volume are excited. The resulting emitted photons are then detected using a wide-field detection system oriented perpendicularly to the light-sheet illumination plane [1,13–16]. This unique imaging configuration provides optical sectioning and ensures that the 2D section inside the excitation volume is imaged in a single shot, dramatically reducing phototoxicity, photobleaching and acquisition times [1,2,15].

A critical condition for achieving high-quality imaging using LSFM is that the illuminating light-sheet plane and the detection focal plane must be co-planar (i.e., parallel and overlapping). However, due to the irregular sample shape, complex inner structures, mismatches of refractive index (RI) and imperfect optical alignment, the co-planar requirement is often violated causing degradations in image quality [1,2]. We previously proposed a deep learning based method to rapidly and accurately perform autofocus during LSFM imaging [17]. This approach enhances the image quality when the illumination plane and the detection plane are parallel (Fig. 1(a)).

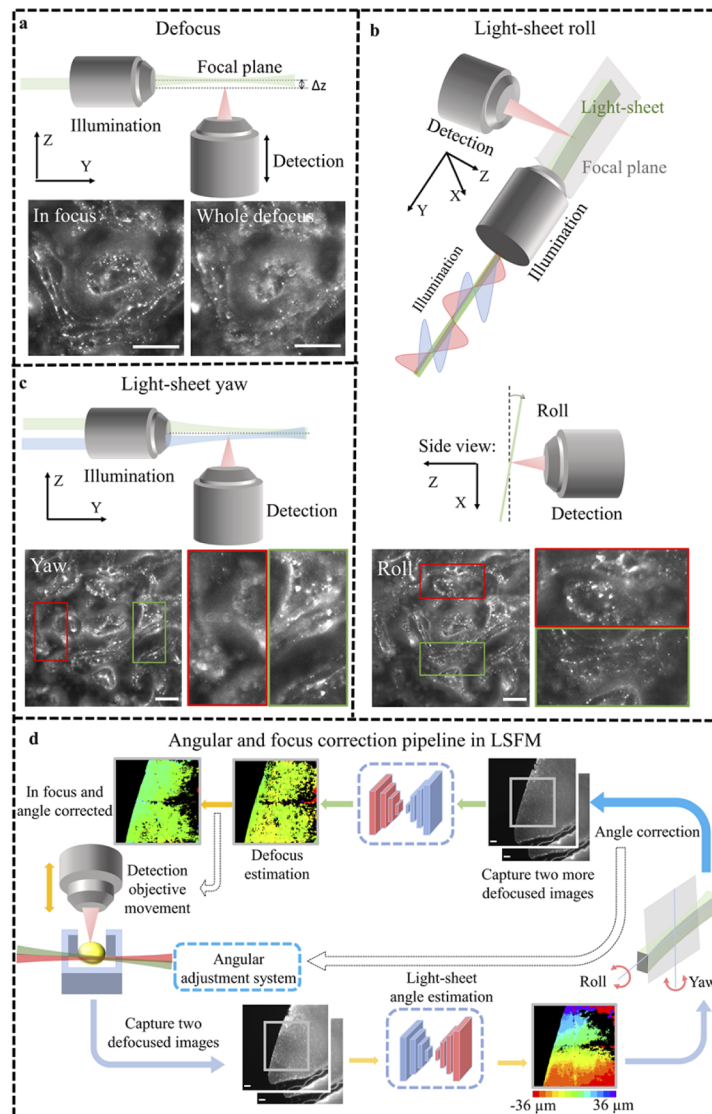


Fig. 1. The effect of the illumination angle on image quality in light-sheet fluorescence microscopy. (a) A motorized stage can move the objective lens and correct for uniform defocus aberration across the field-of-view. (b) and (c) Using two dual galvo scanning systems the light sheet setup can control both the roll and yaw angles of the light-sheet illumination beam, respectively. Sample induced roll or yaw angles in the illumination beam result in variable image quality within the field of view, in focus regions are marked by a green rectangle, while out of focus regions are marked by a red rectangle. Roll (b) and yaw (c) angles will result in aberrations that are vertically and horizontally oriented respectively. (d) Overview of the illumination angle correction pipeline. During image acquisition, two defocused images are acquired and sent to a deep learning network to predict the defocus distance on a pixel level. From the defocus map, the yaw and roll angles of the light-sheet relative to the detection objective focal plane can be calculated and corrected. After angle correction, the image is refocused again using the network and the entire field-of-view is brought in focus. In the color bar, the red and purple colors represent the extreme defocus values of -36 μm and 36 μm respectively. Black color in the processed image represents detected background or pixels with low certainty. Scale bars, 100 μm .

However, refocusing fails to address cases where the illumination beam and the detection focal plane are not parallel. In this case the illumination beam has residual roll (Fig. 1(b)) and yaw (Fig. 1(c)) angles that cause different regions of the field-of-view to experience different defocus levels. Consequently, focusing on a single uniform value will not enhance the image quality in the entire field-of-view. These sample-induced angular errors were previously reported in live imaging of intact embryos [1,2], and we routinely observe these angular errors when imaging tissue cleared bones (e.g., porcine cochlea, and femur) with high resolution (> 0.5 numerical aperture). Bones contain irregular curved components that can refract the illumination beam. Fortunately, if the error in the yaw and roll angles is known, a correction to the illumination beam can be applied in real-time using a spatial-light-modulator or a galvanometric scanner [2].

Manually observing and correcting for these angular errors is highly time-consuming and laborious. For example, an adult pig cochlea consists of dozens of tiles (~ 30 -60) and each tile depth can be more than 5 mm deep, making the manual estimation and correction of sample-induced angular errors for so many sections in advance impractical. To address this issue, autofocus methods [2,18–26] that estimate the defocus level in various regions in the image could be used. From the spatially varying defocus values, the residual yaw and roll angles of the illumination beam can be estimated and corrected. Traditional autofocus methods that acquire and evaluate a stack of images (~ 10 -20) are not ideal for angular error estimation since they are slow. Moreover, traditional autofocus methods provide suboptimal results in the presence of spherical or other non-defocus aberrations that are common in tissue clearing applications [26], and their performance sharply degrades when analyzing relatively small fields-of-view. The latter limitation is especially restrictive, since the defocus level in multiple small regions in the field-of-view is required to estimate the angular error.

To solve this limitation, we utilize deep learning methods recently employed in biomedical imaging [27–34]. These tools are particularly useful in our case since they only require one or two images to estimate the defocus distance, and they perform well on relatively small and aberrated image patches [35]. In general, several studies applied deep learning-based methods to solve the autofocus problem, using one [19,36–38] or two images [39,40]. These fast autofocus methods were applied to tissue slides, and reflection microscopy of semiconductors.

Herein, we extend our previous work [9,12,17] on autofocus LSFM and devised a deep learning-based algorithm to estimate the angle of the light-sheet relative to the detection objective focal plane (Fig. 1(d)). This algorithm is integrated in our LSFM setup and captures two defocused images. The images are fed into a deep learning network to generate a defocus map, which is then used to estimate and correct the angular and defocus aberrations by utilizing two galvo scanners and a linear stage (Fig. 1(d)). We show that our angular correction approach performs well on multiple tissue types (e.g., brain, cochlea and lungs), by successfully overlapping the light-sheet illumination plane with the detection objective focal plane. This new angular correction approach may be used to correct sample induced angular aberrations without human intervention, and greatly improve the image quality in high resolution 3D imaging.

2. Methods

2.1. Sample preparation

In this study, the cochlea and lung samples were extracted from euthanized pigs, while the brain samples were extracted from Mosaic Analysis for Double Marker (MADM) mice [34]. The cochlea samples were cleared and labeled using a modified BoneClear protocol [9,41], while the rest of the samples were cleared and labeled using the iDISCO protocol [42–44]. The cochlea samples were labeled using Myosin VIIa (CY3 as secondary), and the brain samples were labeled against GFP (Alexa Fluor 647 as secondary) and RFP (CY3 as secondary). All animals were harvested under the regulation and approval of the Institutional Animal Care and Use Committee (IACUC) at North Carolina State University.

2.2. Light sheet setup

All the images were captured using a custom-built LSFM that was designed based on previous protocols [1,12,17]. The setup was described in detail in [17], and here we briefly describe the important components that control the illumination beam angle. The light-sheet plane illumination was generated by a continuous-wave laser (Coherent OBIS LS 561-50). The laser beam was expanded and focused on the sample to create a static Gaussian beam. To generate the light-sheet, the Gaussian beam was dithered at a high frequency (500 Hz) using a 2D scanning galvo (Fig. 2(a); Cambridge Technology; 6215H) and a dual-channel arbitrary function generator (Tektronix; AFG31022A). The 2D scanning galvo could also control the roll angle, by synchronizing the frequencies and phases of the voltages, which drive the vertical and horizontal mirrors inside the galvo system. One mirror was dedicated for scanning the light-sheet beam up and down (x-axis), and the other mirror was responsible for the position of the light sheet on the z-axis (see Fig. S1). By adjusting the amplitude of the galvo mirror (in z-axis) the roll angle was determined and controlled in real time. Fig. S1 shows the voltage diagram, which was produced by the arbitrary function generator in order to create a certain roll angle. Again, the phase and frequency of the two channels needed to be locked to produce a stable light-sheet with a well-defined roll angle. To control the yaw and pitch angles, a pivot galvo was used (Thorlabs; GVS202). The pivot angle moved the illumination spot location over the 2D scanning galvo, and therefore on the conjugated back aperture of the illumination lens. The changes in the spot location resulted in angular changes on the sample plane. For the detection path, the perpendicular detection objective lens (10 \times , 0.6 numerical aperture, Olympus, XLPLN10XSVM-P2) together with a tube lens and a CMOS camera (Hamamatsu, C13440-20CU) were used to acquire a wide field image of the sample. The schematic diagram of our LSFM is shown in Fig. 2(a) and other hardware information about our microscope can be found in our recent publications [9,17].

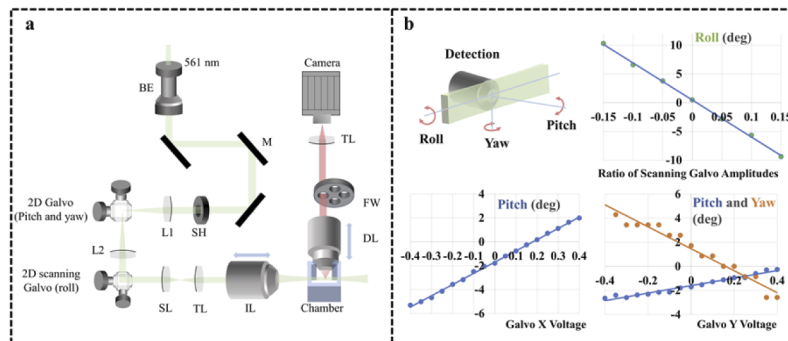


Fig. 2. A schematic diagram of the LSFM and the relationship between galvo control voltages and the light sheet illumination angle. (a) A schematic of the light sheet setup. The green and red paths are the illumination and detection paths respectively. The position and angle of the light-sheet are controlled by two galvo-mirror systems: 2D galvo (pitch and yaw angles), and 2D scanning galvo (roll angle). To adjust the detection focal plane, the detection objective lens is moved by a linear translation stage. BE: beam expander; M: mirror; SH: shutter; L: lens; SL: scan lens; TL: tube lens; IL: illumination lens; DL: detection lens; FW: filter wheel. (b) The linear relationship between galvo control voltage and the pitch, yaw, and roll angles. In the Roll angle, the ratio of scanning galvo means the ratio of voltage amplitudes applied on the two scanning mirrors. Please note the small coupling between the yaw and pitch angles.

To correct the illumination angle, the yaw, pitch and roll angles of the light sheet were measured as a function of galvo voltages. The angles were measured without any specimen in the imaging chamber. First, the emission filter was removed. Second, the 2D scanning galvo voltage was

set to a constant, and a Gaussian beam profile was observed on the camera. Third, the beam waist location and approximate propagation direction in the volume were recorded. Fourth, the voltage that controls either the yaw, pitch and roll angles was changed to a constant value. Last, the new position of the beam waist location and its new propagation direction in the volume were recorded. To find the new position of the beam the detection objective had to be refocused. From the change in the waist location and propagation direction, the roll, pitch and yaw angles were estimated. By repeating this procedure for different voltages, we found a linear relationship between the galvo control voltage and the resulting angles as shown in Fig. 2(b). The linear gradients between the galvo voltage and light sheet angle (Grad_{yaw} and $\text{Grad}_{\text{roll}}$) were later used for online angular correction.

2.3. Datasets acquisition for training and testing

To train and test the deep learning models, we acquired images from our custom-built LSM. Towards this end, the cleared tissue samples were placed in the chamber filled with dibenzyl-ether (DBE). For the training dataset, we first focused the images and set the illumination angles manually in order to achieve a uniform image quality in the field-of-view. Then the detection objective automatically moved and captured a stack of defocused images with defocus values ranging from $-36\ \mu\text{m}$ to $36\ \mu\text{m}$ with $6\ \mu\text{m}$ step size (13 images) [17]. The acquired images were uniformly defocused without any observable angle. Although the models were trained on image stacks that exhibit only parallel but defocused illumination, they generalized well when angular errors were present. This greatly simplified the training process, as we did not have to train the dataset on angled illumination that its absolute angle value was also affected by the specimen in an unpredictable way. A detailed description on how to train the model based on these uniformly acquired stacks of images can be found in [17]. The test set that included cases with tilted illumination required substantial manual labeling.

To test the network's ability to estimate the yaw and roll angles, image stacks with random yaw and roll angles were collected (~ 60 stacks). The images in each stack had a non-uniform image quality because of the introduced angular aberration. Test image stacks were captured in the same method as the training dataset i.e., 13 defocused images. The angle in these stacks was evaluated manually, since we could not rely on the galvo voltage to determine the angular error in the presence of a sample, which induced angular errors. Towards this end, we generated a graphical user interface that facilitated the labeling process (Fig. S2), as the regional defocus levels must be estimated manually at multiple sub regions in the image.

2.4. Network Implementation and training process

For accurately estimating the illumination angle of the light-sheet, it was beneficial to estimate the defocus values in the image at the pixel level, rather than patch level. Therefore, we used a U-net architecture [30] for multi-classification of the defocus levels at the pixel level (Fig. 3). We decided to use a classification network, and not to train the network to infer the defocus value directly, as the datasets generation process was easier. In light sheet microscopy, when focusing an image manually, it was easier to assign the value of the defocus inside a $6\ \mu\text{m}$ bin, instead of trying to estimate it precisely. This is the case since as long as the objective focal plane was approximately within the light sheet FWHM the image remained sharp. Here, we extended the U-net to support two defocus images with an equal distance, which improved the accuracy of the defocus classification. We also added a certainty measure [17,35,45] to the U-net, to reject background and low certainty pixels. The network classified each pixel into 13 classes, whereby each class represented a different defocus distance between the light sheet and the focal plane of the detection objective (Δz_{est}). The Δz_{est} values in the 13 classes were equal to $-36, -30, -24, \dots, 24, 30, 36\ \mu\text{m}$. The $6\ \mu\text{m}$ space between the classes was determined empirically (see [17]) and we believe it is related to the full width half maximum (FWHM) of the light sheet. If the objective

focal plane was approximately within the light sheet FWHM (on average $\sim 14 \mu\text{m}$ across the entire field of view) the image remained sharp.

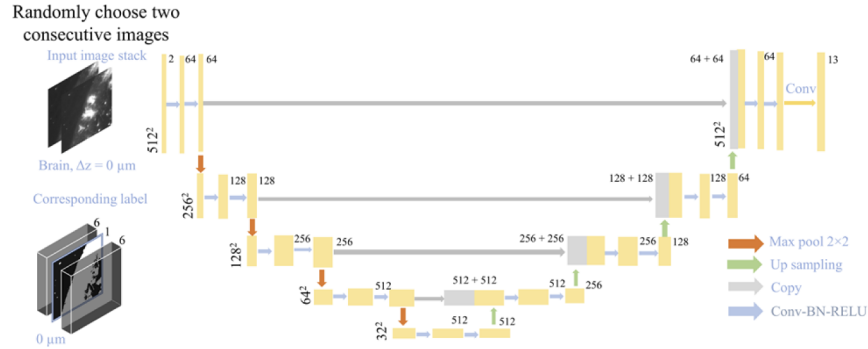


Fig. 3. The architecture of the network. In the example, a tissue cleared brain image with true label of $\Delta z = 0 \mu\text{m}$ is shown. The two images from the training set have $6 \mu\text{m}$ defocus distance between them, and they serve as input to the U-net. The output of the network has the same size as the input but with 13 channels. The value in each channel represents the probability of the defocus distance to be in one of the 13 classes. Each of the 13 classes represents Δz_{est} from $-36 \mu\text{m}$ to $36 \mu\text{m}$ with $6 \mu\text{m}$ step size. The ground truth label is a binary stack of 13 channels, in this case the class that represents $\Delta z = 0 \mu\text{m}$ is equal to 1 and zero elsewhere. BN: batch normalization; RELU: rectified linear unit; Conv: convolutional layer.

The input to the U-net was two defocused images $6 \mu\text{m}$ apart ($I_{\Delta z}$ and $I_{\Delta z + 6 \mu\text{m}}$; 512×512 pixels each; Δz is the real and unknown defocus distance), and we previously found that $6 \mu\text{m}$ step size between the defocused images provided the optimum results in terms of classification accuracy [17]. The U-net output is a matrix ($512 \times 512 \times 13$) with the same number of rows and columns as in the original input images, plus 13 channels. For a given pixel, the values in the 13 channels represented the probability distribution vector of Δz_{est} belonging the 13 defocus classes. To train the network, a dataset containing 379 image stacks was used. The label for each pair of input images was a matrix ($512 \times 512 \times 13$) with the value 1 corresponding to the channel with the true defocus value of the images (Δz). Even in the channel that corresponded to the true defocus value, pixels that corresponded to background (i.e., very low intensity) were assigned to zero. The rest of the channels in the matrix were assigned with zero.

From the network's output ($512 \times 512 \times 13$), a defocus distance prediction map was generated (512×512 pixels) by assigning the defocus value with the highest probability for each pixel.

2.5. Illumination angle correction in LSM

In the illumination angle correction workflow, the network was accessed twice; once to correct for angle, and a second time to correct for defocus (Fig. 1(d)). Two consecutive images $I_{\Delta z}$ and $I_{\Delta z + 6 \mu\text{m}}$ ($6 \mu\text{m}$ defocus distance apart) were captured and fed into the U-net model. From the network output, a defocus distance prediction map was generated ($\Delta z_{est}(x, y)$), and a linear regression was used to fit a plane to the prediction map. The fitted plane parameters provided estimates for the roll and yaw angles ($roll_{predict}$ and $yaw_{predict}$) using the following process: from n random Δz_{est} values, the illumination plane equation can be derived using least square solver

for β_0 , β_1 and β_2 (plane parameters), using the following equation:

$$[\Delta z_{est}(x_1, y_1), \Delta z_{est}(x_2, y_2), \dots, \Delta z_{est}(x_n, y_n)]' = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Using the estimated values of β_0 , β_1 and β_2 the yaw and roll angles can be calculated with the knowledge of the pixel size ($6.5 \times 6.5 \mu\text{m}^2$) and the size of the image. The following equations were used, under the assumption that the yaw and roll angles are small (i.e., less than 7°). For larger angles, the accuracy will improve if the transformation will be derived from the plane parameters using for instance Rodrigues' rotation formula.

$$\Delta z_{est}(x, y) = \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 \end{bmatrix}'$$

$$yaw_{predict} = \text{atan}((\Delta z_{est}(\text{left}, \text{mid}) - \Delta z_{est}(\text{right}, \text{mid})) / ((\text{left} - \text{right}) \times \text{pixel size}))$$

$$roll_{predict} = \text{atan}((\Delta z_{est}(\text{mid}, \text{top}) - \Delta z_{est}(\text{mid}, \text{bot})) / ((\text{top} - \text{bot}) \times \text{pixel size}))$$

Atan is the arctangent function, left, and right, refer to the horizontal edge coordinates of the image (e.g., 1 and 2048, respectively). Top and bot refer to the vertical edge coordinates of the image (e.g., 1 and 2048, respectively), and mid refers to the middle coordinate of the image (e.g., 1024). For correcting the angles, the changes to the galvo voltage were calculated as follows:

$$\begin{cases} \Delta galvo_{voltage} = (-yaw_{predict}) / Grad_{yaw} \\ \Delta roll_{ratio} = (-roll_{predict}) / Grad_{roll} \end{cases}$$

Note that the linear gradients $Grad_{yaw}$ and $Grad_{roll}$ (Fig. 2(b)) between the galvo voltage and light sheet yaw and roll angles were found in section 2.2. Hence, the updated galvo voltages for the galvo scanners were $galvo_{updated} = galvo_{current} + \Delta galvo_{voltage}$ and $roll_{updated} = roll_{current} + \Delta roll_{ratio}$. After the angle correction, the light sheet plane was parallel to the focal plane of the detection objective. Then, an additional two consecutive images were captured for refocusing and fed into the network again. Based on all the pixel level predictions, the most abundant class was considered as the defocus distance ($\Delta z_{predict}$). Accordingly, the detection objective was translated using a motorized linear stage (Newport; CONEX-TRB12CC with SMC100CC motion controller) to refocus the image. All in all, the workflow required four images to correct for the illumination angle and defocus distance.

The control software and graphical user interface (app designer in MATLAB) of the custom-built LSFM were implemented in MATLAB R2019b environment. The trained network can also be accessed in MATLAB environment.

3. Results

3.1. Prediction of the illumination angle and defocus

The network was implemented in Python 3.5 with PyTorch-1.4.0 Deep learning framework and trained on one NVidia Tesla V100-32GB GPU on Amazon Web Service for approximately 12 hours. A binary cross entropy loss function was used between the provided mask and the output of network. The learning rate was set to $1e-5$ with Adam optimizer. Data augmentation methods such as normalization, random crop, saturation and flip were also used.

We tested our U-net model to perform autofocus in comparison with our previous deep neural network (DNN) model [17], and traditional autofocus image methods (DCTS: Shannon entropy of the normalized discrete cosine transform, and TENV: Tenengrad variance). In comparison to the U-net model, the DNN model provided regional defocus values, but not pixel-based predications. Consequently, the DNN model was less suitable to predict errors in the illumination angles [17]. The test dataset included 55 image stacks with uniform defocus values across the field of view (Fig. 4(a)). Image patches with the size of either 128×128 pixels ($83 \times 83 \mu\text{m}^2$) or 512×512 pixels ($333 \times 333 \mu\text{m}^2$) were randomly cropped, and 13 random defocus positions were used from each stack. For the DNN and U-net models, only two defocused images were required, while traditional autofocus methods (TENV and DCTS) were fed with the full image stack as input (13 images). The average absolute distance between the ground truth and the prediction $\Delta z_{\text{predict}}$ were calculated per each of the autofocus methods (Fig. 4(a)). Since the neural network methods (U-net and DNN) provided either pixel or regional level data, $\Delta z_{\text{predict}}$ for the entire image was defined as the most abundant defocus prediction.

For small image patches (128×128 pixels), U-net, DNN, and DCTS achieved an average distance error of 6.82, 7.14, and 8.56 μm , and for larger image patch (512×512 pixels), the average distance errors were 5.62, 6.02, and 5.86 μm , respectively (Fig. 4(a)). Overall, our developed U-net model performed better or comparable with the DNN and traditional methods on different image sizes. For a detailed confusion matrix of the U-net model see Fig. S3. Interestingly, traditional autofocus methods like DCTS had larger average distance error when they performed on small image patches. Hence, they were less suitable for angular predictions, which require the division of the image into small image patches. Therefore, deep learning-based methods that provide pixel-based defocus measurements were preferable for illumination angle estimation.

To test the angular predication of the workflow (Fig. 4(b)), we manually observed and labeled 60 image stacks with different yaw and roll angles. To calculate the angles manually, the defocus distance in various positions in the image was calculated, and a plane was fitted to these measurements. From the plane equation, the ground truth yaw and roll angles were derived. Then, the test images were fed to the network, and the difference between the ground truth angles and the U-net output results was calculated (Fig. 4(b)). The relative average angle error of U-net for the roll and yaw angles were 0.53° and 0.63° respectively. While for the DCTS, the relative average angle error of roll and yaw were 0.59° and 1.30° respectively. For side by side comparison between DCTS and Unet in terms of performance and speed see Table S1. We believe that the asymmetry in the error (i.e., roll versus yaw) was related to the fact that the light sheet width was broadening at the edges of the field-of-view. This fact made it difficult to accurately determine the in-focus position at the edges with great accuracy (see discussion).

Figure 4(c) shows the U-net prediction maps of four representative test images, with relatively uniform defocus. The color bar indicates the defocus distance from the focal plane. The U-net average output ($\Delta z_{\text{predict}}$) matched the ground truth defocus distance well. Figure 4(d) shows the prediction maps in cases where the light-sheet illumination was tilted relative to the detection objective. The prediction maps showed vertical and horizontal color gradient. The gradient features of the prediction maps demonstrated the U-net ability to generalize well on the non-uniform defocused images, although it was trained on uniformly defocused images. The angular prediction allowed to correct for sample induced angular errors in the illumination beam.

3.2. Online deep learning-based illumination correction pipeline improved image quality in the porcine cochlea and mouse brain

We integrated the illumination correction pipeline with our custom-built LSFM and then we performed perturbation experiments on tissue cleared pig cochlea (Fig. 5(a)) and mouse brain (Fig. 5(b)). In these experiments, the light sheet galvo control variables were randomly chosen, and the distorted images are presented in Fig. 5(a1) and 5(b1). The original image size was

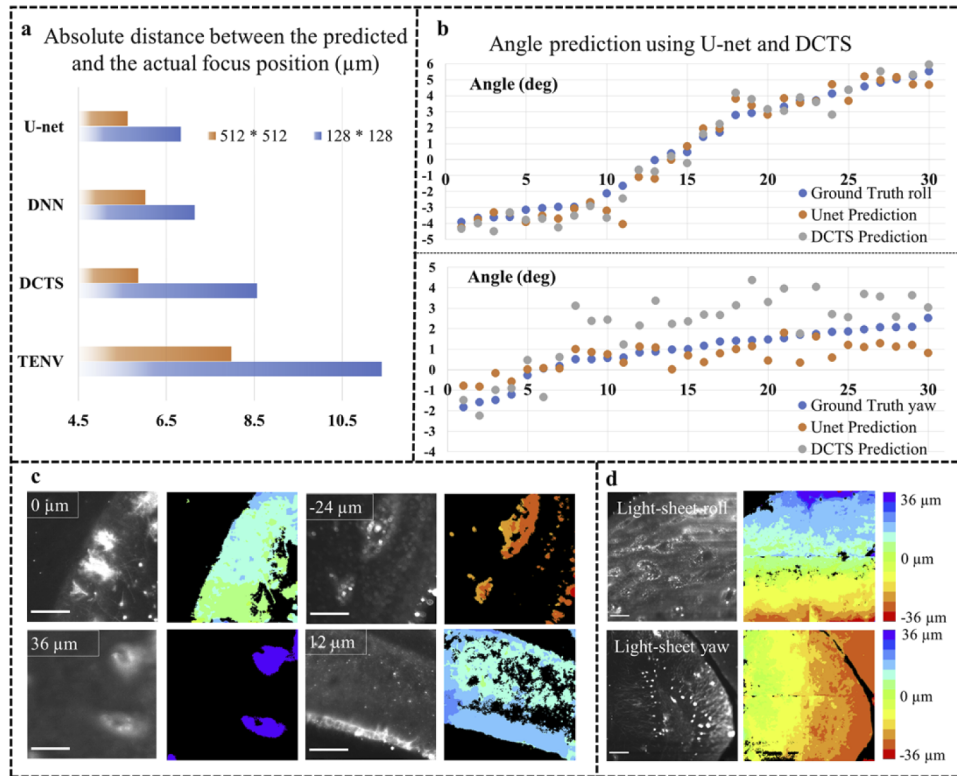


Fig. 4. Pixel level estimation of defocus and illumination angle. (a) Performance comparison between the U-net model, deep neural network (DNN) and traditional autofocus methods (DCTS and TENV) over 550 random cases. The U-net and DNN received as input only two defocused images while the traditional methods received a full stack of images (13 images). U-net, DNN and DCTS achieve comparable performance on large image patches (512×512 pixels), while on small image patches (128×128), U-net and DNN outperform DCTS and TENV. (b) The angle prediction of U-net and DCTS over 30 roll-tilted and 30 yaw-tilted images, the ground truth labels were calculated manually. (c) Representative examples of defocus distance prediction on the defocus dataset (image size: 512×512). The color bar indicates the relative distance of light sheet to the focal plane (prediction maps). (d) The U-net prediction maps for images that were captured using tilted light-sheet illumination. The images size is 1024×1024 pixels, and the prediction maps (right column) are stitched by 4 smaller patches (512×512). Scale bars, $100 \mu\text{m}$.

2048×2048 pixels, and to generate the defocus prediction maps, the middle part was cropped (1024×1024 pixels). The gradients in the prediction maps were evident and accordingly the yaw and roll angles were calculated. After the illumination angle was corrected, the gradients in the prediction maps were less noticeable, and most areas in the prediction maps shared the same color (Fig. 5(a2) and 5(b2)). This indicated that the illumination and detection planes were approximately parallel. Then the angle corrected images were sent through the same network again to refocus the image after angle correction. The most abundant color in the prediction map was extracted and the associated defocus distance $\Delta z_{\text{predict}}$ was inferred. Then the detection lens was refocused on the illumination plane ($\Delta z_{\text{predict}}$ distance) and the image quality improved dramatically in the field-of-view (Fig. 5(a3) and 5(b3)). After the correction, we ran the network again to verify that the illumination plane and the detection focal plane were indeed co-planer.

This can be observed as the prediction maps in Fig. 5(a3) and 5(b3) has a greenish color (0 μm defocus).

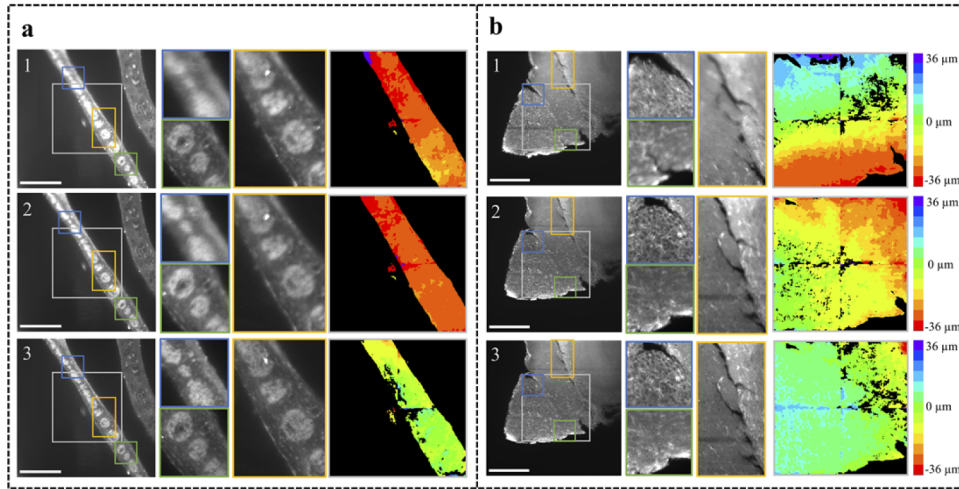


Fig. 5. Online perturbation experiments in LSFM. (a1 and b1) Light-sheet images of a cochlea and brain tissue, respectively. In these images, the illumination plane and the detection plane are not parallel or focused. **(a2 and b2)** Images of the same area after illumination angle correction. Notice how the color gradients are less observable in the prediction maps. **(a3 and b3)** Images of the same area after defocus and illumination angle correction. The blue, green and yellow boxes mark the position of the zoom-in images, and the gray box mark the area that was used to generate the prediction maps. The improved image quality in a3 and b3 shows the pipeline ability to make the light sheet plane and the detection focal plane co-planar. Note that the image content may differ because the light sheet spatial position changed when the angle was corrected, and therefore slightly different areas in the sample are illuminated and recorded. Scale bars, 400 μm .

3.3. Performance of online deep learning-based illumination correction pipeline on unseen-tissue

We then tested our workflow on unseen tissue i.e., tissue cleared lung sample that was extracted from a pig. Please note that the U-net training stage was not conducted on any lung tissue, and this stage was performed to evaluate the network's ability to generalize to other untrained tissue types. In general, the lung tissue had very different morphological features than the cochlea and the brain, as can be seen in Fig. 6(a1) and 6(b1), that were acquired with tilted illumination. Figure 6(a2) and 6(b2) are the same images as in Fig. 6(a1) and 6(b1) but after illumination angle correction. After correcting for the defocus (Fig. 6(a3) and 6(b3)) the overall image quality improved. As expected, we noticed a performance drop on an unseen tissue, since the lung features were different from the brain and cochlea. Also, the pig lung tissue was not filled with agarose during the tissue processing and therefore, it collapsed and was denser than a mouse lung that can be agarose inflated [46].

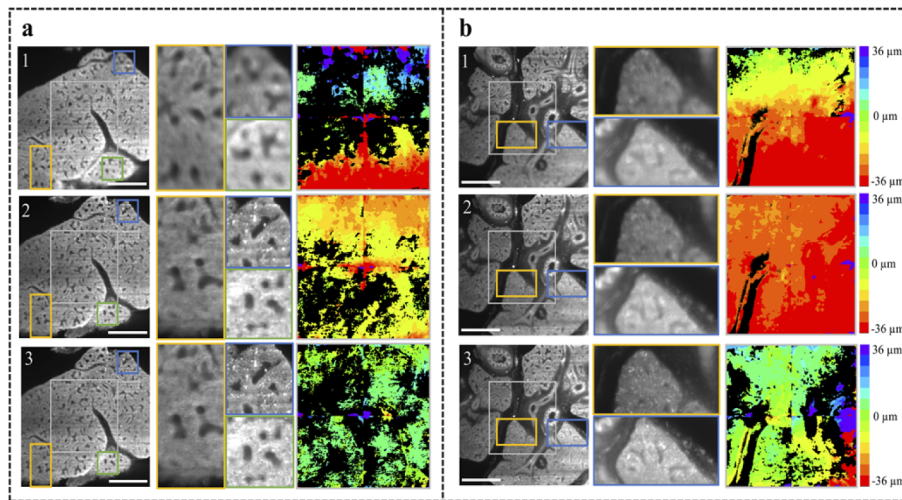


Fig. 6. Perturbation experiments on unseen lung tissue. (a1 and b1) Light-sheet images of a tissue cleared porcine lung sample. In these images, the illumination plane and the detection plane are not parallel or focused. **(a2 and b2)** Light-sheet images after angle correction. The color gradients are reduced. **(a3 and b3)** Images of the same area after defocus correction. The overall improvement in the image quality is evident. The blue, and yellow boxes mark the position of the zoom-in images, and the gray box mark the area that was used to generate the prediction maps. Scale bars, 400 μm .

4. Discussion and conclusions

For high resolution LSFM imaging, image quality is greatly degraded if the light-sheet illumination is not parallel to the objective detection plane. This angular component in the light-sheet illumination can result either from inadequate alignment of the optical system, or from refraction of the illumination beam inside the specimen. Here, we develop an online method to correct for both the angle and defocus in LSFM using a deep learning model. Our model estimates the defocus distance on a pixel level using only two defocused images at a fixed step size. Based on the pixel level information on the defocus distances in the image, the angular error of the light sheet illumination is estimated. The degrees of freedom in our adaptive light-sheet microscope are then used to correct for the uneven defocus over the entire image. We describe a procedure to calibrate the light-sheet angle as function of galvo scanner voltage, so users can replicate the method in independent platforms. In comparison with traditional autofocus methods, our new approach is faster since only two images are required to estimate the defocus level versus a full image stack. Additionally, the new approach is more accurate, as traditional autofocus methods yield less precise outputs when working on small image patches that are required for estimation of the angles. We tested the utility of our method, and we observe significant improvement in image quality in mouse brain, as well as porcine cochleae and lung samples. This approach can lighten the burden of human intervention in acquiring 3D datasets without manual calibration steps and improve image quality overall.

Simple and alternative approaches to mitigate the effects of non-uniform illumination are to either increase the diameter of the illumination beam used to generate the light sheet, or to use a low numerical aperture objective lens. A larger diameter beam increases the illuminated volume in the sample, which increases the chances for the detection focal plane to overlap with the illuminated volume in the sample, even in the presence of an angle between the two. However, increasing the illumination beam diameter diminishes the optical sectioning, which results in

low quality images. Alternatively, one can use low numerical aperture objectives (e.g., 1–2 \times , numerical aperture <0.1), whose depth of field will be large enough to capture moderately tilted light-sheet illumination beams. Again, this solution comes at the expense of low axial and spatial resolution.

Although successful in its original goal, our new approach possesses a few limitations. First, we find that the yaw angle of the light-sheet is more difficult to estimate than the roll angle. We believe that the primary reason for this limitation is that the Gaussian beam diverges at the edges relative to the center of the field-of-view where it is tightly focused. In this case, the optical sectioning is compromised, making it more difficult to observe significant changes in the image quality across multiple defocus levels in the edge of the field-of-view. As such, the estimation for the yaw angle is more difficult. In the future, we can incorporate “non-diffractive” beams (i.e., the beam spread out less while propagating away from the focal plane, e.g., Bessel beam). Such an application may improve the estimation of the yaw angle. Second, the main current computational bottleneck that restricts real-time correction is the processing time of the certainty measure. In general, the camera captures an image with 2048×2048 pixels, however we only use the middle part (1024×1024 pixels) as an input to the network for two reasons: to accelerate the processing time, and since the FWHM of the light sheet is broader at the edges, hence, the network’s performance degrades. Therefore, we only train the network on the middle part. Given a patch of 1024×1024 pixels, both the focus correction and the angle correction take approximately 4 seconds, without including the acquisition time. Approximately 2 seconds are required to generate the prediction map for 1024×1024 pixels in Python. While it takes the network only a few milliseconds (<0.1 seconds) to generate the defocus distance per pixel, the calculation of the certainty requires ~ 1 s. In future efforts we will generate a less computationally intensive certainty measure, which may provide similar results to our current approach. The reduction in the processing time should bring the proposed method from online to real-time.

Funding. Life Sciences Research Foundation; National Institutes of Health (R01NS098370, R56NS117019).

Acknowledgments. The authors would like to thank Dr. Jorge Piedrahita and the NCSU Central Procedure Lab for their help with tissue collection.

Disclosures. The authors declare no conflicts of interests.

Data availability. The source code of training deep learning model in this paper is available at GitHub [47]. The image dataset underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. L. A. Royer, W. C. Lemon, R. K. Chhetri, and P. J. Keller, “A practical guide to adaptive light-sheet microscopy,” *Nat. Protoc.* **13**(11), 2462–2500 (2018).
2. L. A. Royer, W. C. Lemon, R. K. Chhetri, Y. Wan, M. Coleman, E. W. Myers, and P. J. Keller, “Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms,” *Nat. Biotechnol.* **34**(12), 1267–1278 (2016).
3. T. Chakraborty, M. K. Driscoll, E. Jeffery, M. M. Murphy, P. Roudot, B.-J. Chang, S. Vora, W. M. Wong, C. D. Nielson, H. Zhang, V. Zhemkov, C. Hiremath, E. D. De La Cruz, Y. Yating, I. Bezprozvanny, H. Zhao, R. Tomer, R. Heintzmann, J. P. Meeks, D. K. Marciano, S. J. Morrison, G. Danuser, K. M. Dean, and R. Fiolka, “Light-sheet microscopy of cleared tissues with isotropic, subcellular resolution,” *Nat. Methods* **16**(11), 1109–1113 (2019).
4. F. Amat, W. Lemon, D. P. Mossing, K. McDole, Y. Wan, K. Branson, E. W. Myers, and P. J. Keller, “Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data,” *Nat. Methods* **11**(9), 951–958 (2014).
5. P. J. Keller, A. D. Schmidt, A. Santella, K. Khairy, Z. Bao, J. Wittbrodt, and E. H. K. Stelzer, “Fast high-contrast imaging of animal development with scanned light sheet-based structured illumination microscopy,” *Nat. Methods* **7**(8), 637–642 (2010).
6. R. Tomer, K. Khairy, F. Amat, and P. J. Keller, “Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy,” *Nat. Methods* **9**(7), 755–763 (2012).
7. P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. K. Stelzer, “Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy,” *Science* **322**(5904), 1065–1069 (2008).

8. Y. Wu, P. Wawrzusin, J. Senseney, R. S. Fischer, R. Christensen, A. Santella, A. G. York, P. W. Winter, C. M. Waterman, Z. Bao, D. A. Colón-Ramos, M. McAuliffe, and H. Shroff, "Spatially isotropic four-dimensional imaging with dual-view plane illumination microscopy," *Nat. Biotechnol.* **31**(11), 1032–1038 (2013).
9. A. Moatti, Y. Cai, C. Li, T. Sattler, L. Edwards, J. Piedrahita, F. S. Ligler, and A. Greenbaum, "Three-dimensional imaging of intact porcine cochlea using tissue clearing and custom-built light-sheet microscopy," *Biomed. Opt. Express* **11**(11), 6181–6196 (2020).
10. Z. Huang, P. Gu, D. Kuang, P. Mi, and X. Feng, "Dynamic imaging of zebrafish heart with multi-planar light sheet microscopy," *J. Biophotonics*, **14**(5), e202000466 (2021).
11. J. N. Singh, T. M. Nowlin, G. J. Seedorf, S. H. Abman, and D. P. Shepherd, "Quantifying three-dimensional rodent retina vascular development using optical tissue clearing and light-sheet microscopy," *J. Biomed. Opt.* **22**(07), 1 (2017).
12. A. Greenbaum, K. Y. Chan, T. Dobrev, D. Brown, D. H. Balani, R. Boyce, H. M. Kronenberg, H. J. McBride, and V. Gradinaru, "Bone CLARITY: Clearing, imaging, and computational analysis of osteoprogenitors within intact bone marrow," *Sci. Transl. Med.* **9**, 387 (2017).
13. P. A. Santi, "Light sheet fluorescence microscopy: a review," *J. Histochem. Cytochem.* **59**(2), 129–138 (2011).
14. M. Weber and J. Huisken, "Light sheet microscopy for real-time developmental biology," *Curr. Opin. Genet. Dev.* **21**(5), 566–572 (2011).
15. E. M. C. Hillman, V. Voleti, W. Li, and H. Yu, "Light-sheet microscopy in neuroscience," *Annu. Rev. Neurosci.* **42**(1), 295–313 (2019).
16. P. J. Keller and H.-U. Dodt, "Light sheet microscopy of living or cleared specimens," *Curr. Opin. Neurobiol.* **22**(1), 138–143 (2012).
17. C. Li, A. Moatti, X. Zhang, H. T. Ghashghaei, and A. Greenbaum, "Deep learning-based autofocus method enhances image quality in light-sheet fluorescence microscopy," *Biomed. Opt. Express* **12**(8), 5214–5226 (2021).
18. F. Shen, L. Hodgson, and K. Hahn, "Digital autofocus methods for automated microscopy," *Methods in Enzymol.*, **414**, 620–632 (2006).
19. Y. Luo, L. Huang, Y. Rivenson, and A. Ozcan, "Single-shot autofocus of microscopy images using deep learning," *ACS Photonics* **8**(2), 625–638 (2021).
20. Y. Liao, Y. Xiong, and Y. Yang, "An auto-focus method of microscope for the surface structure of transparent materials under transmission illumination," *Sensors* **21**(7), 2487 (2021).
21. J. Liao, Y. Jiang, Z. Bian, B. Mahrou, A. Nambiar, A. W. Magsam, K. Guo, S. Wang, Y. ku Cho, and G. Zheng, "Rapid focus map surveying for whole slide imaging with continuous sample motion," *Opt. Lett.* **42**(17), 3379–3382 (2017).
22. M. Bathe-Peters, P. Annibale, and M. J. Lohse, "All-optical microscope autofocus based on an electrically tunable lens and a totally internally reflected IR laser," *Opt. Express* **26**(3), 2359–2368 (2018).
23. K. Guo, J. Liao, Z. Bian, X. Heng, and G. Zheng, "InstantScope: a low-cost whole slide imaging system with instant focal plane detection," *Biomed. Opt. Express* **6**(9), 3210–3216 (2015).
24. L. Silvestri, M. C. Müllenbroich, I. Costantini, A. P. Di Giovanna, G. Mazzamuto, A. Franceschini, D. Kutra, A. Kreshuk, C. Checcucci, L. O. Toresano, P. Frascioni, L. Sacconi, and F. S. Pavone, "Universal autofocus for quantitative volumetric microscopy of whole mouse brains," *Nat. Methods* **18**(8), 953–958 (2021).
25. L. Silvestri, M. C. Müllenbroich, I. Costantini, A. P. D. Giovanna, L. Sacconi, and F. S. Pavone, "RAPID: Real-time image-based autofocus for all wide-field optical microscopy systems," *bioRxiv*, 170555 (2017).
26. Y. Tian, K. Shieh, and C. F. Wildsoet, "Performance of focus measures in the presence of nondefocus aberrations," *J. Opt. Soc. Am. A* **24**(12), B165–B173 (2007).
27. S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* (2016), pp. 514–517.
28. C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.* **63**(13), 135007 (2018).
29. T. D. Silva, E. Y. Chew, N. Hotaling, N. Hotaling, C. A. Cukras, and C. A. Cukras, "Deep-learning based multi-modal retinal image registration for the longitudinal analysis of patients with age-related macular degeneration," *Biomed. Opt. Express* **12**(1), 619–636 (2021).
30. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *MICCAI* (2015).
31. H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: a full-scale connected unet for medical image segmentation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2020), pp. 1055–1059.
32. P. Rajpurkar, J. Irvin, R. L. Ball, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. P. Langlotz, B. N. Patel, K. W. Yeom, K. Shpanskaya, F. G. Blankenberg, J. Seekins, T. J. Amrhein, D. A. Mong, S. S. Halabi, E. J. Zucker, A. Y. Ng, and M. P. Lungren, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLoS Med.* **15**(11), e1002686 (2018).
33. Q. Que, Z. Tang, R. Wang, Z. Zeng, J. Wang, M. Chua, T. S. Gee, X. Yang, and B. Veeravalli, "CardioXNet: automated detection for cardiomegaly based on deep learning," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2018), pp. 612–615.

34. Y. Cai, X. Zhang, S. Z. Kovalsky, H. T. Ghashghaei, and A. Greenbaum, "Detection and classification of neurons and glial cells in the MADM mouse brain using RetinaNet," *PLoS One* **16**(9), e0257426 (2021).
35. S. J. Yang, M. Berndt, D. Michael Ando, M. Barch, A. Narayanaswamy, E. Christiansen, S. Hoyer, C. Roat, J. Hung, C. T. Rueden, A. Shankar, S. Finkbeiner, and P. Nelson, "Assessing microscope image focus quality with deep learning," *BMC Bioinformatics* **19**(1), 77 (2018).
36. Y. Wu, Y. Rivenson, H. Wang, Y. Luo, E. Ben-David, L. A. Bentolila, C. Pritz, and A. Ozcan, "Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning," *Nat. Methods* **16**(12), 1323–1331 (2019).
37. H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, and L. Waller, "Deep learning for single-shot autofocus microscopy," *Optica* **6**(6), 794–797 (2019).
38. S. Jiang, J. Liao, Z. Bian, K. Guo, Y. Zhang, and G. Zheng, "Transform- and multi-domain deep learning for single-frame rapid autofocusing in whole slide imaging," *Biomed. Opt. Express* **9**(4), 1601–1612 (2018).
39. T. Ivanov, A. Kumar, D. Sharoukhov, F. Ortega, and M. Putman, "DeepFocus: a deep learning model for focusing microscope systems," in *Applications of Machine Learning 2020* (International Society for Optics and Photonics, 2020), 11511, p. 1151103.
40. T. R. Dastidar and R. Ethirajan, "Whole slide imaging system using deep learning-based automated focusing," *Biomed. Opt. Express* **11**(1), 480–491 (2020).
41. K. H. R. Jensen and R. W. Berg, "Advances and perspectives in tissue clearing using CLARITY," *J. Chem. Neuroanat.* **86**, 19–34 (2017).
42. N. Renier, Z. Wu, D. J. Simon, J. Yang, P. Ariel, and M. Tessier-Lavigne, "iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging," *Cell* **159**(4), 896–910 (2014).
43. T. Liebmann, N. Renier, K. Bettayeb, P. Greengard, M. Tessier-Lavigne, and M. Flajolet, "Three-dimensional study of Alzheimer's disease hallmarks using the iDISCO clearing method," *Cell Rep.* **16**(4), 1138–1152 (2016).
44. D. T. Mzinza, H. Fleige, K. Laarmann, S. Willenzon, J. Ristenpart, J. Spanier, G. Sutter, U. Kalinke, P. Valentin-Weigand, and R. Förster, "Application of light sheet microscopy for qualitative and quantitative analysis of bronchus-associated lymphoid tissue in mice," *Cell. Mol. Immunol.* **15**(10), 875–887 (2018).
45. C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.* **27**(3), 379–423 (1948).
46. R. J. E. van den Bijgaart, N. Kong, C. Maynard, and V. Plaks, "Ex vivo live imaging of lung metastasis and their microenvironment," *J. Visualized Exp.* (108), e53741 (2016).
47. C. Li, "Source code for angle correction network," (1), Github, (2021).
https://github.com/Chenli235/AngleCorrection_Unet.